Extremal Quantum Correlations and Cryptographic Security

T. Franz,[1,*] F. Furrer,[1] and R. F. Werner[1]

[1] *Institut für Theoretische Physik, Leibniz Universität Hannover*
Appelstraße 2, 30167 Hannover, Germany

We investigate a fundamental property of device independent security in quantum cryptography by characterizing probability distributions which are necessarily independent of the measurement results of any eavesdropper. We show that probability distributions that are secure in this sense are exactly the extremal quantum probability distributions. This allows us to give a characterization of security in algebraic terms. We apply the method to common examples for two-party as well as multi-party setups and present a scheme for verifying security of probability distributions with two parties, two measurement settings, and two outcomes.

The idea of using quantum systems for secure communication has been around for more than 25 years now. But still the boundaries of quantum cryptography have not been fully understood. Only recently a remarkable feature of quantum systems has been realized, namely that observed violations of a Bell inequality may imply cryptographic security, even if the measurements that lead to the violation are unknown to legitimate parties. This principle goes under the name "device independent security" and has been proven against collective attacks [1], and recently against arbitrary attacks for memoryless measurement devices [2, 3]. But still no proof for the most general situation is known. In this paper we focus on the question, when measurement outcomes obtained by the legitimate parties are independent of measurements performed by an eavesdropper. We give a necessary and sufficient condition for this under the assumption that the probability distributions are known without error.

We consider a quantum correlation experiment with N separated parties, each performing one of M different local measurements with K outcomes. We denote this situation by the triple (N, M, K). In a device independent scenario the parties (usually N=2) want to extract a secret key from the observed correlations in which the security estimation is solely based on the measured probability distributions. There are no assumption on the proper functioning of the measurement devices or the measured system, e.g., on their dimension. Probability distributions that are useful for cryptography have to feature certain properties. First, the obtained correlations should not permit a local hidden variable (LHV) model as in this case a potential adversary can have full knowledge about the correlations. Second, the correlations should only be weakly correlated to any possible measurement of an adversary. The first property is well known to be equivalent to violate a Bell inequality (see below), but the latter still lacks a concrete characterization.

In this paper we address this problem by specifying all probability distributions which do not allow a LHV model and are provably statistically independent of the knowledge of any eavesdropper. We show that these probability distributions, which we call secure, can completely be characterized in geometric terms. Indeed in the convex body $\mathcal{Q}$ of all quantum probability distributions the secure points are precisely the non-classical extremal points, i.e., those which are not deterministic and cannot be obtained as proper convex combination of other points in $\mathcal{Q}$.

The characterization of extremal points in $\mathcal{Q}$ is of general interest and numerical approaches to determine them are known [4, 5]. In our examples, we provide and discuss different tools to certify, respectively, find extremal probability distributions for particular (N, M, K)-cases. In many situations it turns out to be easier to establish a stronger property, i.e., that the algebraic structure of the measurement operators is completely determined by the probability distributions. This also leads to a stronger notion of security. The most prominent example (see example 3) are correlations which maximally violate the Clauser-Horne-Shimony-Holt (CHSH)-inequality [6].

Our results have links to previous results obtained in the framework of non-signaling correlations, i.e., theories that are more general than quantum theory. One direction of our result, namely that extremality implies security was proven in [7] for non-signalling theories in the bipartite case. In this paper, we only discuss the quantum framework, although our proofs can in principle be adapted to any non-signaling theory.

Definitions. For simplicity, we consider the general (N, M, K) case, even though the results are also valid for different numbers of measurement settings and outcomes for each party. We denote the probability for obtaining a string of outcomes $\underline{x} = (x_1, ..., x_N)$ given a string of measurement settings $\underline{s} = (s_1, ..., s_N)$ by $\mathbb{P}(\underline{x}|\underline{s})$. These numbers are assumed to be known exactly, i.e., we do not consider the uncertainties involved in estimating such probabilities from a finite sample.

The set of probability distributions $\mathbb{P}$ conform to a LHV model, i.e., which can be realized by assuming the measurements reveal outcomes whose probabilities are predetermined, is called the set $\mathcal{C}$ of classical correlations. It is a polytope, i.e. generated by a finite number

*Electronic address: torsten.franz@itp.uni-hannover.de

of extremal points, which are given by the assignment of definite outcomes to each measurement. The faces (of maximal dimension) correspond to inequalities, which are linear in $\mathbb{P}$, and are called (tight) Bell inequalities. In the $(2,2,2)$ case all tight Bell inequalities are equivalent to the CHSH inequality [8]. A survey about Bell inequalities and further references can be found in [9]

We are interested in the set $\mathcal{Q}$ of quantum correlations, which is defined as the set of all probability distributions $\mathbb{P}$ that can be realized by a quantum representation

$$\mathbb{P}(\underline{x}|\underline{s}) = \mathrm{tr}(\rho F(\underline{x}|\underline{s})), \tag{1}$$

where ρ is a density operator on a Hilbert space $\mathscr{H}$, whose dimension is not constrained and can be infinite, and $F(\underline{x}|\underline{s}) = F_1(x_1|s_1)\cdots F_N(x_N|s_N)$ is a product of commuting operators on $\mathscr{H}$. $\{F_i(x|s)\}_x$ are the measurement operators of the observable chosen by the i^{th} party according to the measurement setting s. Thus the $F_i(x|s)$ are positive operators satisfying $\sum_{x=1}^{K} F_i(x|s) = \mathbb{1}$, and have to commute for different sites, since the parties are independent. As shown in Appendix A, every $\mathbb{P}$ which can be realized in this way can also be realized in a simplified "standard" form, in which $\rho = |\Omega\rangle\langle\Omega|$ is a pure state and the operators $F_i(x|s)$ are projections. Moreover, in the standard form $|\Omega\rangle$ is cyclic for the algebra $\mathcal{A}(F)$, which is obtained from the $F_i(x|s)$ by taking products, linear combinations and limits of expectation values. Cyclic means that the vectors $A|\Omega\rangle$ with $A \in \mathcal{A}(F)$ span a dense subspace in $\mathscr{H}$.

The set $\mathcal{Q}$ is a closed convex set which has in contrast to $\mathcal{C}$ a continuum of extremal points (see for instance [10]). Bell inequalities define the boundary between $\mathcal{C}$ and $\mathcal{Q}$. The set $\mathcal{Q}$ can be characterized similarly by inequalities that are linear in $\mathbb{P}$, satisfied by all $\mathbb{P} \in \mathcal{Q}$ and tight for at least one $\mathbb{P} \in \mathcal{Q}$. We call them Tsirelson inequalities. For every linear expression in $\mathbb{P}$ there is a maximum on $\mathcal{C}$, and another, usually larger one on $\mathcal{Q}$ which leads to a Tsirelson inequality. Computational methods to derive such maximal violations in $\mathcal{Q}$ are derived in [4, 5]. For the CHSH expression in the $(2,2,2)$-case these maxima are 2 [6] and $2\sqrt{2}$ [11], respectively. The value 4 is achieved on the set of "no-signalling correlations" $\mathcal{P}$, defined by the property that the measurement of one party does not change the probabilities observed by another. Similar to $\mathcal{C}$, $\mathcal{P}$ is generated by finitely many extremal points [7]. It holds with proper inclusion $\mathcal{C} \subset \mathcal{Q} \subset \mathcal{P}$.

Secure probability distributions. We model the eavesdropper by another quantum party, whose measurements must commute with all $F(\underline{x}|\underline{s})$. Accordingly, we call a probability distribution $\mathbb{P}$ *secure*, if $\mathbb{P}$ does not factorize, i.e., $\mathbb{P}(\underline{x}|\underline{s}) \neq \prod_{j=1}^{N} \mathbb{P}_j(x_j|s_j)$, and for any quantum representation and any operator E commuting with all $F_i(x_i|s_i)$

$$\mathrm{tr}\Big(\rho E F(\underline{x}|\underline{s})\Big) = \mathrm{tr}(\rho E)\,\mathbb{P}(\underline{x}|\underline{s}). \tag{2}$$

The operator E represents all possible measurements an eavesdropper could perform. The requirement that $\mathbb{P}$ is not a product is necessary to exclude classical deterministic points, i.e. the extremal points of $\mathcal{C}$, for which (2) is satisfied trivially. As we will see, this excludes all probability distributions which can be realized in LHV models.

In device independent cryptography, our definition ensures that an attack of an eavesdropper can never be better than a classical guess. The number of extractable secure bits by classical postprocessing can then be characterized by the classical smooth min-entropy [12].

Our first main result gives a geometric interpretation of secure probability distributions: *A probability distribution $\mathbb{P}$ is secure, if and only if it is extremal in $\mathcal{Q}\backslash\mathcal{C}$.*

The argument is straightforward. Suppose, $\mathbb{P}$ is secure, but not extremal. Then there exists a direct sum representation and a convex decomposition with $\mathbb{P} = \lambda\mathbb{P}_1 + (1-\lambda)\mathbb{P}_2$, $0 \leq \lambda \leq 1$. Now use the definition (2) with E being the projector onto the first/second summand to get $\mathbb{P} = \mathbb{P}_1$ and $\mathbb{P} = \mathbb{P}_2$. This shows that the convex combination is indeed trivial and $\mathbb{P}$ is extremal. As all extremal correlations in $\mathcal{C}$ are of product form, it follows that $\mathbb{P} \notin \mathcal{C}$. Conversely, suppose $\mathbb{P}$ is extremal and $\mathbb{P} \notin \mathcal{C}$. As before, we can conclude that $\mathbb{P}$ cannot be of product form. Take any commuting $0 < E < \mathbb{1}$ and set $\lambda = \mathrm{tr}(\rho E)$. Define $\mathbb{P}_1 = (1/\lambda)\mathrm{tr}(\rho E F(\underline{x}|\underline{s}))$ and $\mathbb{P}_2 = (1/(1-\lambda))\mathrm{tr}(\rho(\mathbb{1}-E)F(\underline{x}|\underline{s}))$ such that $\mathbb{P} = \lambda\mathbb{P}_1 + (1-\lambda)\mathbb{P}_2$. As $\mathbb{P}$ is extremal, it holds that $\mathbb{P} = \mathbb{P}_1$, which is just equation (2), so $\mathbb{P}$ is secure.

To decide whether a given probability distribution is secure has now been reduced to certifying extremality in $\mathcal{Q}$. This is in general a hard problem. Even in the $(2,2,2)$-case no simple algebraic constraints are known to verify extremality of a given $\mathbb{P}$. In this paper, we will provide an explicit, yet limited certification scheme in example 3.

Algebraically secure probability distributions. There is a straightforward way to strengthen the definition of secure probability distributions by extending the factorization property to a larger set of observables. The reason is that the stronger notion of security is often easier to verify.

A probability distribution $\mathbb{P}$ is called *algebraically secure*, if it is secure and for any quantum representation and any operator E commuting with all $F_i(x_i|s_i)$

$$\mathrm{tr}\Big(\rho E\tilde{F}\Big) = \mathrm{tr}(\rho E)\,\mathrm{tr}(\rho\tilde{F}), \tag{3}$$

for all $\tilde{F} \in \mathcal{A}(F)$.

They are characterized as follows: *A probability distribution $\mathbb{P}$ is algebraically secure, if and only if it is extremal in $\mathcal{Q}\backslash\mathcal{C}$ and has a unique quantum representation, up to unitary equivalence.*

A sketch of the proof goes as follows. Assume first that $\mathbb{P}$ is algebraically secure, and therefore extremal. Let $\rho = |\Omega\rangle\langle\Omega|$ together with $F_i(x_i|s_i)$, and $\rho' = |\Omega'\rangle\langle\Omega'|$ with $F'_i(x_i|s_i)$ be two representations of $\mathbb{P}$ on suitable Hilbert spaces $\mathscr{H}$, $\mathscr{H}'$. Condition (3) implies that for all corresponding operators $\tilde{F} \in \mathcal{A}(F)$ and $\tilde{F}' \in \mathcal{A}(F')$,

$\mathrm{tr}(\rho\tilde{F}) = \mathrm{tr}(\rho'\tilde{F}')$. Otherwise, the direct sum representation with E chosen as the projector on the first or second summand contradicts (3). Define then the unitary operator U via $U\tilde{F}|\Omega\rangle = \tilde{F}'|\Omega'\rangle$ which transforms one representation into the other. Because $|\Omega\rangle$ and $|\Omega'\rangle$ are cyclic U can be extended to a unitary from $\mathscr{H}$ to $\mathscr{H}'$. Conversely, assume that $\mathbb{P}$ is extremal and all representations are unitarily equivalent. Let $0 \leq E \leq \mathbb{1}$ be an operator commuting with all $F_i(x_i|s_i)$. Since $\mathbb{P}$ is extremal, $\frac{1}{\mathrm{tr}(\rho E)}\sqrt{E}\rho\sqrt{E}$ together with the operators $F_i(x_i|s_i)$ is a valid quantum representation of $\mathbb{P}$. Hence, $E = \mathbb{1}$, which implies (3).

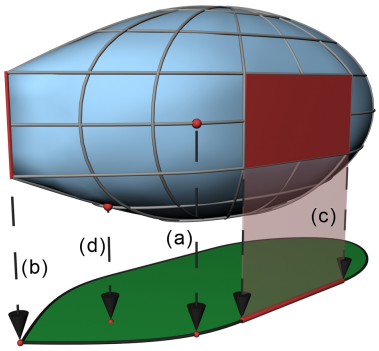

FIG. 1: Sketch of the set of quantum representations $\mathcal{S}$ (above) and the set of probability distributions $\mathcal{Q}$ (below). An extremal probability distribution can either correspond to a unique point (a) or to a face of $\mathcal{S}$ (b). Other faces of $\mathcal{S}$ can be mapped to faces of $\mathcal{Q}$ (c). Not all extremal points of $\mathcal{S}$ are also extremal for $\mathcal{Q}$ (d).

Secure vs. algebraically secure. It is now interesting to identify cases for which the notions of secure and algebraically secure coincide. To formalize the question, we can introduce a map Γ from all possible (unitary inequivalent) quantum representations $\mathcal{S}$ $(= \mathcal{S}(N, M, K))$ to the set of probability distributions $\mathcal{Q}$. The set $\mathcal{S}$ can be considered as a convex set and the map Γ is linear and surjective, but not injective. The extremal points of $\mathcal{S}$ are exactly the irreducible quantum representations, which are defined by the property that the only invariant subspaces of $\mathcal{A}(F)$ are $\{0\}$ and $\mathscr{H}$. As shown in [13], each extremal probability distribution $\mathbb{P} \in \mathcal{Q}$ admits an irreducible quantum representation. Hence, a secure probability distribution $\mathbb{P}$ is algebraically secure if and only if $\Gamma^{-1}(\mathbb{P})$ is exactly one extremal point in $\mathcal{S}$. In FIG. 1, the point (a) corresponds to an algebraically secure probability distribution, while the point (b) and the endpoints of the line (c) are secure, but not algebraically secure.

In the following we discuss examples, for which we provide methods to find extremal points and criterions to decide when they are also algebraically secure.

Example 1: The (N,2,2)-case. The algebraic structure of the $(N, 2, 2)$-case is quite well understood (see e.g. [14] and references therein). All irreducible quantum representations are in this case given on an N-qubit subspace $\mathscr{H} = \otimes_{i=1}^{N}\mathbb{C}^2$ with an arbitrary pure state $\psi \in \mathscr{H}$ and measurements, which are parameterized by N angles $\theta_1, ..., \theta_N$ $(\theta_i \in [0, \pi])$. The measurements are given at site i as $F_i(1, 1) = \frac{1}{2}(\mathbb{1} + \sigma_3)$ and $F_i(1, 2) = \frac{1}{2}(\mathbb{1} + \sin(\theta_i)\sigma_1 + \cos(\theta_i)\sigma_3)$, together with their complements $F_i(2, s) = \mathbb{1} - F_i(1, s)$. The σ_i denote the Pauli matrices and we omitted the identities on the tensor factors for the other parties. This parametrization in $\{\theta_i\}$ and ψ is sufficient to determine the whole convex body $\mathcal{Q}$. An arbitrary $\mathbb{P}$ is a direct sum of at most $4^N + 1$ irreducible representations. Compare [15] for an alternative deviation of these results.

In order to find extremal points and test algebraic uniqueness we combine the above parametrization with a maximization of a Tsirelson inequality. More explicitly, for each functional given by coefficients $\{c(\underline{x}|\underline{s})\}$, we can ask for the maximal quantum violation, i.e., $Q_c := \sup_{\mathbb{P}\in\mathcal{Q}} \sum_{\underline{x},\underline{s}} c(\underline{x}|\underline{s})\mathbb{P}(\underline{x}|\underline{s})$. In general, Q_c can be computed by a hierarchy of semi-definite programs [4, 5]. Here, we follow another strategy by parameterizing the corresponding operator $C = \sum c(\underline{x}|\underline{s})F(\underline{x}|\underline{s}) = C(\theta_1, ..., \theta_N)$ by means of the irreducible representations. The maximization of $\langle\psi|C(\theta_1, ..., \theta_N)|\psi\rangle$ over all $\theta_i \in [0, \pi)$ and $\psi \in \mathbb{C}^{2^N}$ yields Q_c. Moreover, if there is exactly one set of parameters $\theta_1, ..., \theta_N$ and a unique state ψ for which the maximum is attained, the corresponding probability distribution $\mathbb{P}$ is algebraically secure. In the case where more than one possible choice of $\theta_1, ..., \theta_N, \psi$ leads to a maximal violation, we can determine the convex span of the corresponding probability distributions. This corresponds to the face given by the intersection of $\mathcal{Q}$ and the hyperplane $\{\mathbb{P} \mid \sum_{\underline{x},\underline{s}} c(\underline{x}|\underline{s})\mathbb{P}(\underline{x}|\underline{s}) = Q_c\}$. Extremal points of that face are extremal points of $\mathcal{Q}$, and thus, secure probability distributions.

As a straightforward application, one can deduce that the probability distributions leading to maximal violation of Mermin's inequalities [16] are algebraically secure.

Example 2: Certificate of extremality in the (2,2,2)-case. The idea of the foregoing example was to find extremal $\mathbb{P}$'s by maximizing a given Tsirelson expression. Here, we start with a particular $\mathbb{P}$ and want to construct a Tsirelson inequality saturated by $\mathbb{P}$. If there exists such an inequality which is not trivial, i.e. cannot be saturated by any LHV model, and no other probability distribution in $\mathcal{Q}$ saturates it (or alternatively that just one quantum representation of $\mathbb{P}$ exists), extremality of $\mathbb{P}$ is certified.

We focus on the (2,2,2)-case and discuss a method how to construct a maximally violated Tsirelson expressions for a given $\mathbb{P}$. It comes along with a natural order of complexity for which we solve the lowest order explicitly. The main ingredient is again the parametrization of the irreducible quantum representations by a state $\psi \in \mathbb{C}^2 \otimes \mathbb{C}^2$ and two angles $\underline{\theta} = (\theta_A, \theta_B)$ (see previous example) for which we denote the obtained probability distribution by $\mathbb{P}_{(\underline{\theta},\psi)}$. Because we are only interested in extremal $\mathbb{P}$s, it is sufficient to consider $\mathbb{P}_{(\underline{\theta},\psi)}$ with a real ψ (see Appendix B). Since we have dichotomic measurements

we can equivalently work with ± 1 valued observables instead of measurement operators. We denote the observables on Alice's (Bob's) side by A_1, A_2 (B_1, B_2) and set $A_0 = B_0 = \mathbb{1}$.

Finding a Tsirelson inequality for $\mathbb{P}_{(\underline{\theta},\psi)}$ is equivalent to the following task: Construct a positive operator $T = \sum_k P_k(A_i, B_j)^\dagger P_k(A_i, B_j)$, with $P_k(A_i, B_j)$ polynomials in $A_i \otimes B_j$, $i, j = 0, 1, 2$, such that (i) $P_k(A_i(\theta_A), B_j(\theta_B))\psi_0 = 0$ for all k and (ii) $T = \sum_{i,j=0}^{2} t_{ij} A_i \otimes B_j$ for all possible observables in $\mathscr{H}$. Here, $A_i(\theta_A)$, $B_j(\theta_B)$ denote the observables of the representation $(\underline{\theta}, \psi)$. Condition (ii) implies that T can be interpreted as a linear functional of $\mathbb{P}$, (i) that it is 0 for $\mathbb{P}_{(\underline{\theta},\psi)}$, and the ansatz for T that T is a positive operator and thus its associated functional on $\mathbb{P}$ is positive for each $\mathbb{P} \in \mathcal{Q}$.

In order to solve the problem a constraint on the degree of the polynomials P_k in the ansatz for T has to be imposed. This introduces a natural hierarchy, where the order limits the possible $\mathbb{P}_{(\underline{\theta},\psi)}$ for which the method succeeds. For the simplest ansatz, $P_k = \sum_{j=1}^{2}(\alpha_{kj} A_j \otimes \mathbb{1} - \beta_{kj} \mathbb{1} \otimes B_j)$ ($\alpha_{kj}, \beta_{kj} \in \mathbb{R}$), the $\mathbb{P}$ for which a Tsirelson inequality can be constructed are exactly the ones which correspond to a representation $(\phi_x^{\pm}, \theta_A, \theta_B)$ with maximally entangled state $\phi_x^{\pm} = \frac{1}{\sqrt{2}}(\cos x, \mp \sin x, \sin x, \pm \cos x)$ ($x \in [0, \pi)$) for which

$$\frac{\sin(2x)\sin(2x \pm \theta_B)}{\sin(2x - \theta_A)\sin(2x - \theta_A \pm \theta_B)} < 0$$

holds. The corresponding Tsirelson inequality and the derivation can be found in Appendix B.

Example 3: The (2,M,2)-case for full correlations. The difficulty of finding extremal points in the $(2, M, 2)$ scenario can be considerably reduced, as it is sufficient to consider only full correlations. This was shown by Tsirelson in [11] where he characterized all extremal points. In the following let A_i, B_j, $i, j \in \{1, \ldots, M\}$, denote ± 1 valued observables located by Alice and Bob, and ρ a density operator. The set of quantum correlations $\mathcal{Q}_{cor}$ is given by all correlation tables $c_{ij} = \mathrm{tr}(A_i B_j \rho)$ which can be obtained by means of a quantum representation. In [11] it was proven that all quantum representations of an extremal correlation table which is not deterministic have uniform marginal distributions $\mathrm{tr}(A_i\rho) = \mathrm{tr}(B_j\rho) = 0$. Thus, non-deterministic extremal correlations in $\mathcal{Q}_{cor}$ correspond to secure probability distributions in $\mathcal{Q}$. Furthermore, an extremal correlation table which allows just one quantum representation gives rise to an algebraically secure point.

For every correlation table c_{ij} exists a so-called *c-system*, that is, a collection of vectors x_i, y_j ($i, j \in \{1, \ldots, M\}$) with $||x_i|| \leq 1$, $||y_j|| \leq 1$ in an Euclidian space with dimension M, such that $c_{ij} = \langle x_i, y_j \rangle$. If $\mathbb{P}$ is extremal, the corresponding c-systems are isometric to each other, $||x_i|| = ||y_j|| = 1$ and the linear hull of the $\{x_i\}$ and $\{y_j\}$ coincide. Calling the dimension of the linear hull the *rank* r of the c-system, it further follows that $\{x_i \otimes x_i, y_j \otimes y_j\}$ span the symmetric subspace of $\mathbb{R}^r \otimes \mathbb{R}^r$. The following inequalities hold: $r \leq M$, $r \leq -1/2 + \sqrt{1/4 + 4M})$ and $r(r+1)/2 \leq 2M - 1$. There are two cases to be distinguished. For c-systems with even rank, the representation is unique (up to unitary equivalence), while for c-systems with odd rank, there are exactly two non-equivalent representations.

With this, the question of secure versus algebraically secure is equivalent to determining the rank of the c-system which corresponds to the given correlation table. According to the inequalities above, it follows directly that all probability distributions in the $(2, 2, 2)$ and $(2, 3, 2)$-case which correspond to non-classical extremal correlations in $\mathcal{Q}_{cor}$ are algebraically secure.

Acknowledgements T.F. acknowledges support from the DFG under grant WE-1240/12-1. F.F. acknowledges support from the Graduiertenkolleg 1463 of the Leibniz Universität Hannover. R.F.W. acknowledges the support of the EU FP7 project COQUIT (contract number 233747).

Appendix A: Standard Form of a Quantum Representation

As described in the paper, we consider a general probability distribution obtained by N parties, having M measurements with K outcomes each. Let $x \in \{1, .., K\}$ denote one local measurement outcome, $s \in \{1, ..., M\}$ one setting and $\underline{x}$ ($\underline{s}$) denote the N-element strings of outcomes (measurements) for all parties. We define a probability distribution $\mathbb{P}$ to be quantum and thus lying in $\mathcal{Q}$, if there exists a Hilbert space $\mathscr{H}$ together with a state ρ and positive operator valued measures $\{F_i(x, s)\}_{x=1}^{K}$, $i = 1, ..., N$ and $s = 1, ..., M$, such that $[F_i(x, s), F_j(x', s')] = 0$ for every $i \neq j$, and it holds that

$$\mathbb{P}(\underline{x}|\underline{s}) = \mathrm{tr}(\rho F(\underline{x}|\underline{s})), \tag{A1}$$

where $F(\underline{x}, \underline{s}) = \prod_{i=1}^{N} F_i(x_i|s_i)$.

The goal is to show that for any such $\mathbb{P}$ one can find a quantum representation in standard form, that is, a representation which consists of projective measurements and a pure cyclic state for $\mathcal{A}(F)$. The proof is given by an explicit construction.

First, we recapitulate the definition of $\mathcal{A}(F)$ and what it means that a vector is cyclic for $\mathcal{A}(F)$. The algebra generated by the operators $F_i(x, s)$, $\mathcal{A}(F)$, is defined as the closure of the set of all linear combinations of products of $F_i(x, s)$, i.e. $\mathrm{span}_{\mathbb{C}}\{\prod_{l=1}^{m} F_{i_l}(x_{i_l}|s_{i_l})|m \in \mathbb{N}\} \subset \mathcal{B}(\mathscr{H})$ with respect to taking expectation values (i.e., the weak* closure). This means that we add every $G \in \mathcal{B}(\mathscr{H})$ for which a sequence $\{G_j\}$ in $\mathrm{span}_{\mathbb{C}}\{\prod_{l=1}^{m} F_{i_l}(x_{i_l}|s_{i_l})|m \in \mathbb{N}\}$ exists, such that $\lim_{k\to\infty} \mathrm{tr}(\rho G_j) = \mathrm{tr}(\rho G)$ for every state ρ in $\mathscr{H}$. In mathematical terminology this is called a von Neumann algebra [17]. We call now a pure state $|\Omega\rangle$ cyclic for $\mathcal{A}(F)$, if the closure of $\{G|\Omega\rangle | G \in \mathcal{A}(F)\}$ is the entire Hilbert space $\mathscr{H}$.

We begin by turning the measurement operators $F_i(x|s)$ into projective ones, by applying a version of the Naimark dilation successively to each observable $F_i(\cdot, s)$. It suffices to do this for one of the observables, provided we verify that in this construction not only the required commutativity conditions are preserved, but also the projection valuedness of any of the other measurements. So in order to turn the observable $F_i(\cdot, s)$ in to a projective measurement, we define the Hilbert space $\widehat{\mathscr{H}} = \bigoplus_{x=1}^{K} \mathscr{H}_x$, where each of the $\mathscr{H}_x$ is a copy of the given Hilbert space $\mathscr{H}$. We denote by P_x the projection onto the summand with label x, and introduce the isometry

$$V : \mathscr{H} \to \widehat{\mathscr{H}} \qquad V\phi = \bigoplus_x \sqrt{F_i(x,s)}\phi.$$

Then we will set $\widehat{F}_i(x,s) = P_x$, so that $V^*\widehat{F}_i(x,s)V = F_i(x,s)$. For other observables at the same site, e.g., $F_i(\cdot, r)$ with $r \neq s$, we set

$$\widehat{F}_i(x,r) = \begin{cases} VF_i(1,r)V^* + (\mathbb{1} - VV^*) & \text{for } x = 1 \\ VF_i(x,r)V^* & \text{for } x > 1 \end{cases}$$

Because V is an isometry, we again have $V^*\widehat{F}_i(x,r)V = F_i(x,r)$ for all x. With V^* we denote the adjoint operator of V. Moreover, $\widehat{F}_i(x,r)^2 = VF_i(x,r)^2V^*$ for $x > 1$ and $\widehat{F}_i(1,r)^2 = VF_i(1,r)^2V^* + (\mathbb{1} - VV^*)$, so that a projective measurement remains projective. For observables at all other sites $j \neq i$ we take $\widehat{F}_j(x,r) = \bigoplus_{x'} F_j(x,r)$, i.e., as the original observable acting the same on each of the summands. Once again, this preserves projective valuedness, and not only satisfies $V^*\widehat{F}_j(x,r)V = F_j(x,r)$, but even the stronger relation $\widehat{F}_j(x,r)V = VF_j(x,r)$. With this relation it is easy to see that the $\widehat{F}_j(x,r)$ for different j (possibly $= i$) commute, so we can form the product $\widehat{F}(\underline{x}|\underline{s})$ unambiguously, and that $V^*\widehat{F}(\underline{x}|\underline{s})V = F(\underline{x}|\underline{s})$. Hence if we define the state $\hat{\rho} = V\rho V^*$, we obtain a quantum representation of the same point $\mathbb{P} \in \mathcal{Q}$, with $\widehat{F}_i(\cdot, s)$ projective measurements.

In order to turn ρ into a pure and cyclic state we can do the Gelfand-Naimark-Segal (GNS) construction (Theorem 2.3.16 in [17]) of the algbra $\mathcal{A}(F)$ with respect to the state ρ. We consider $\mathcal{A}(F)$ together with the positive semidefinite sesquilinear form defined by $\langle A|B\rangle = \mathrm{tr}(\rho A^*B)$ as a pre-Hilbert space. To get a Hilbert space we first take the quotient with respect to the left ideal $I = \{A \in \mathcal{A}(F) | \mathrm{tr}(\rho A^*A) = 0\}$ and then the completion with respect to the scalar product $\langle\cdot|\cdot\rangle$. We denote the obtained Hilbert space by $\widehat{\mathscr{H}}$ and its elements (in the densely defined subspace) are given by the equivalence classes $\psi_A = \{\tilde{A} | \tilde{A} = A + J, J \in I\}$ for $A \in \mathcal{A}(F)$. We define the representation π of $\mathcal{A}(F)$ on $\widehat{\mathscr{H}}$ by the equation $\pi(A)\psi_B = \psi_{AB}$ for $A, B \in \mathcal{A}(F)$. This representation is a *-homomorphism, that is, it respects products and the adjoint operation. Hence, the operators $\widehat{F}_i(x|s) = \pi\big(F_i(x|s)\big)$ satisfy the same commutation relation as $F_i(x|s)$ and furthermore, projections are mapped onto projections. If we set $\Omega = \psi_{\mathbb{1}}$, we have that $\langle\Omega|\widehat{F}(\underline{x}|\underline{s})\Omega\rangle = \mathrm{tr}(\rho F(\underline{x}|\underline{s})) = \mathbb{P}(\underline{x}|\underline{s})$. We therefore found a quantum representation of $\mathbb{P}$ given by $\widehat{F}_i(x|s)$ and a pure state $|\Omega\rangle$ which is by definition cyclic.

Appendix B: Lowest Order of the (2,2,2)-Certificate

The goal is to check extremality for a given $\mathbb{P}$ in the (2,2,2)-case. We use the same notation as introduced in the example 2 in the paper. Since we are only interested in extremal $\mathbb{P}$ in $\mathcal{Q}$, we can restrict to the ones which belong to an irreducible quantum representation (see example 1 in the paper). They are described in a Hilbert space $\mathscr{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ and parameterized by a state $\psi \in \mathscr{H}$ and angles $\theta_A, \theta_B \in [0, \pi]$, which specify the ± 1-valued observables A_1, A_2 and B_1, B_2 on Alice's and Bob's side. The concrete form of the observables are given by $A_i(\theta_A) = \sum_j t(\theta_A)_{ij} X_j$ and $B_i(\theta_B) = \sum_j t(\theta_B)_{ij} X_j$ with $X_1 = \sigma_1$, $X_2 = \sigma_3$ and

$$t(\theta) = \begin{pmatrix} 0 & 1 \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Because the observables are all real, an extremal $\mathbb{P}$ always allows a representation with a real ψ. To see this, we can write $\psi = \phi + i\eta$ with ϕ, η real vectors in $\mathscr{H}$. Note that if $(\psi, \theta_A, \theta_B)$ generates $\mathbb{P}$ so does $(\bar{\psi}, \theta_A, \theta_B)$, with $\bar{\psi}$ the complex conjugate of ψ. Hence, the state $\rho = \frac{1}{2}(|\psi\rangle\langle\psi| + |\bar{\psi}\rangle\langle\bar{\psi}|) = \frac{1}{2}(|\phi\rangle\langle\phi| + |\eta\rangle\langle\eta|)$ together with θ_A, θ_B generates the same $\mathbb{P}$. But if $\mathbb{P}$ is extremal then also $(\phi, \theta_A, \theta_B)$ and $(\eta, \theta_A, \theta_B)$ generates $\mathbb{P}$, thus, the state can be chosen to be real.

Moreover, the case $\sin\theta = 0$ corresponds to the case where the observables at Alice's or Bob's side commute, which corresponds to $\mathbb{P}$ which can be generated by a LHV model. Hence, we restrict our attention to representations with a real ψ and $\theta \neq 0, \pi$.

We want to construct $T = \sum_{i=1}^{2} P_i(A_k, B_l)^\dagger P_i(A_k, B_l)$ with

$$P_i = \sum_{j=1}^{2} (\alpha_{ij} A_j \otimes \mathbb{1} - \beta_{ij} \mathbb{1} \otimes B_j) \tag{B1}$$

where α and β are matrices in $M_2(\mathbb{C})$, such that the conditions

(i) $P_i(A_l(\theta_A), B_j(\theta_B))\psi = 0$ for $i = 1, 2$

(ii) $T = \sum_{i,j=0}^{2} c_{ij} A_i \otimes B_j$ for all possible observables A_i, B_j in $\mathscr{H}$

are satisfied. Note that a possible observable A_i has to satisfy $A_i^* = A_i$ and $A_i^2 = \mathbb{1}$.

The restricted form of P_i limits the possible irreducible representation for which the method applies, which means that it is not always possible to find coefficients α, β such that condition (i) and (ii) are satisfied. The goal is to determine for which representations

this can be done and derive the corresponding Tsirelson inequality.

We start by analyzing condition (i). Using the particular form of the observables $A_i(\theta_A)$ and $B_j(\theta_B)$ expressed through $t(\theta)$, we find that $P_i\psi = 0$, $i = 1, 2$, results in

$$[X_i \otimes \mathbb{1}]\psi = \sum_j \eta_{ij}[\mathbb{1} \otimes X_j]\psi \quad (i = 1, 2) \tag{B2}$$

where $\eta = t(\theta_A)^{-1}\alpha^{-1}\beta t(\theta_B)$. We assumed here that α is invertible. However, this is not a restriction since otherwise the state ψ is of product form.

In the following it is convenient to use the isomorphism between $\mathscr{H}$ and the Hilbert space $M_2(\mathbb{C})$ with the Hilbert-Schmidt inner product. States $\phi = (\phi_1, \phi_2, \phi_3, \phi_4)$ in $\mathscr{H}$ are identified with matrices

$$\hat{\phi} = \begin{pmatrix} \phi_1 & \phi_2 \\ \phi_3 & \phi_4 \end{pmatrix}$$

and $[A \otimes \mathbb{1}]\phi$ (resp. $[\mathbb{1} \otimes B]\phi$) can be written as $A\hat{\phi}$ (resp. $\hat{\phi}B^T$). Moreover, we have that ϕ is a purification of the density matrix $\rho = (\hat{\phi}^*\hat{\phi})^T$ on $\mathbb{C}^2$. Equation (B2) is then equivalent to

$$X_i\hat{\psi} = \sum_j \eta_{ij}\hat{\psi}X_j. \tag{B3}$$

The following assertion characterizes condition (i): *ψ admits an η such that (B2) is satisfied if and only if $\hat{\psi}^T\psi \propto \mathbb{1}$. Then, it holds that*

$$\eta_{ij} = \frac{1}{2}\mathrm{tr}(\hat{\psi}^{-1}X_i\hat{\psi}X_j). \tag{B4}$$

The proof goes as follows. First, we note that $\hat{\psi}$ must be invertible. This is due to the fact that otherwise the reduced state of ψ given by $(\hat{\psi}^*\hat{\psi})^T$ has determinant 0 and is therefore a pure state. We then multiply equation (B3) with $X_k\hat{\psi}^{-1}$ from the left to find $\sum_j \eta_{ij}X_kX_j = X_k\hat{\psi}^{-1}X_i\hat{\psi}$. Recalling that $\mathrm{tr}(X_kX_j) = 2\delta_{kj}$, we can take the trace and obtain (B4).

We turn now to the first part of the statement. Multiplication from the right of (B3) with $\hat{\psi}^{-1}$ shows that $X_i = \sum_j \eta_{ij}\hat{\psi}X_j\hat{\psi}^{-1}$. Thus, we obtain that

$$\mathrm{tr}(X_iX_k) = \sum_{j,l} \eta_{ij}\eta_{kl}\mathrm{tr}(X_jX_l),$$

from which follows that $\eta\eta^T = \mathbb{1}$. On the other hand one can check that the set G of all $\hat{\psi}$ for which there exists a η such that (B2) holds and $\det(\hat{\psi}) = 1$, describes a group together with the usual matrix multiplication. Moreover, the map $\hat{\psi} \mapsto \eta(\hat{\psi})$ induced by (B4) is a group homomorphism such that $\eta(\hat{\psi}^T) = \eta^{-1}$. From this we can then conclude that $\hat{\psi}^T\hat{\psi} \propto \mathbb{1}$ is the necessary and sufficient condition to solve (B2).

Because condition (i) is satisfied if and only if $\frac{1}{\det\hat{\psi}}\hat{\psi}$ is an orthogonal matrix, the possible states ψ are parameterized by

$$\phi_x^{\pm} = \frac{1}{\sqrt{2}}(\cos x, \mp\sin x, \sin x, \pm\cos x) \tag{B5}$$

where $x \in [0, \pi)$. The state ψ determines the corresponding η uniquely through equation (B4).

Since the reduced state of ψ is equal to $(\hat{\psi}^*\hat{\psi})^T$, it follows directly that $\phi_x^{\pm}$ is maximally entangled. From this follows also that the expectation values of all local observables A_l and B_j vanish.

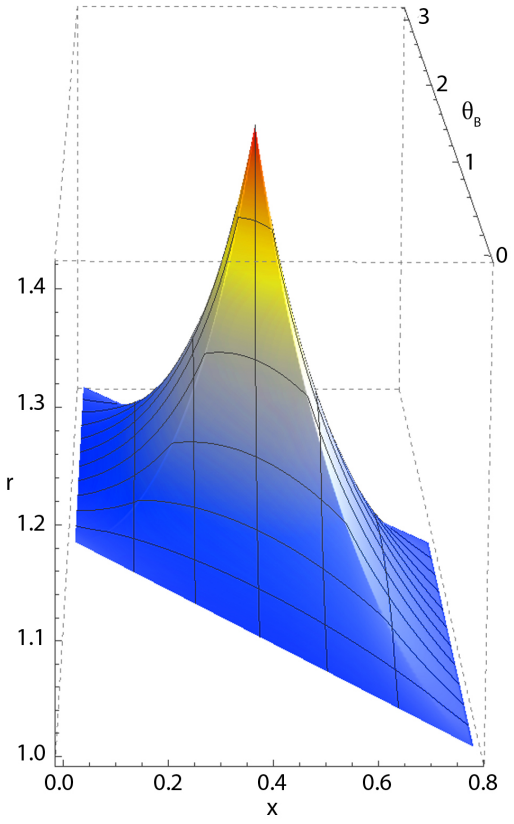

FIG. 2: The plot shows the ratio r between the maximal quantum and classical violation of the Tsirelson inequality (B6) corresponding to $(\psi = \phi_x^+, \theta_A = \frac{\pi}{2}, \theta_B)$ for $\theta_B \in (0, \pi)$ and $x \in (0, \pi/4)$. The domain is restricted due to condition (B7). The quotient r exhibits a $\pi/4$-periodic behavior in x and the peak of $\sqrt{2}$ at $x = \pi/8$ and $\theta_B = \pi/2$ corresponds to a CHSH inequality.

We turn now to condition (ii) and compute the expectation value of T with respect to $\phi \in \mathscr{H}$,

$$\begin{aligned}\langle T\rangle_\phi =& \mathrm{tr}(\alpha^*\alpha + \beta^*\beta) \\ &- \sum_{j,k}(\alpha^*\beta + (\beta^*\alpha)^T)_{jk}\langle A_j \otimes B_k\rangle_\phi \\ &+ \sum_{j\neq k}((\alpha^*\alpha)_{jk}\langle A_jA_k\rangle_\phi + (\beta^*\beta)_{jk}\langle B_jB_k\rangle_\phi).\end{aligned}$$

Thus, condition (ii) requires that the matrices $\alpha^*\alpha$ and $\beta^*\beta$ are diagonal. In this case the Tsirelson inequality reads

$$\sum_{j,k}(\alpha^*\beta+(\beta^*\alpha)^T)_{jk}\langle A_j\otimes B_k\rangle\leq \mathrm{tr}(\alpha^*\alpha+\beta^*\beta). \quad \text{(B6)}$$

The coefficients c_{ij} in condition (ii) are therefore $c_{jk}=(\alpha^*\beta+(\beta^*\alpha)^T)_{jk}$ for $j,k=1,2$, $c_{00}=-\mathrm{tr}(\alpha^*\alpha+\beta^*\beta)$, and the others 0.

Since the expectation value of T is invariant under scaling and simultaneous unitary transformation of α and β, we can without loss of generality assume that $\alpha=\mathrm{diag}(1,\lambda)$ with $\lambda>0$. This can always be achieved by the polar decomposition. Using now that $\eta=t(\theta_A)^{-1}\alpha^{-1}\beta t(\theta_B)$ we can write $\beta=\alpha\gamma$ with $\gamma=t(\theta_A)\eta t(\theta_B)^{-1}$. The condition that $\beta^*\beta$ is diagonal is then equivalent to

$$\lambda^2=-\frac{\overline{\gamma_{11}}\gamma_{12}}{\overline{\gamma_{21}}\gamma_{22}}>0, \quad \text{(B7)}$$

which completely characterizes condition (ii).

Let us now summarize the results from the discussions of the two conditions. Condition (i) says that the only states ψ for which the method applies are of the form (B5). Hence, we can constrain to representations (ψ,θ_A,θ_B) with $\psi=\phi_x^\pm$. For these states we can compute η via equation (B4), and insert it into $\gamma=t(\theta_A)\eta t(\theta_B)^{-1}$ to find that

$$\gamma_x^\pm=\tfrac{1}{\sin\theta_B}\begin{pmatrix}\pm\sin(2x\pm\theta_B) & \mp\sin(2x)\\ \pm\sin(2x-\theta_A\pm\theta_B) & \mp\sin(2x-\theta_A)\end{pmatrix}$$

Condition (B7) can then be computed to be

$$(\lambda_x^\pm)^2=-\frac{\sin(2x)\sin(2x\pm\theta_B)}{\sin(2x-\theta_A)\sin(2x-\theta_A\pm\theta_B)}>0.$$

Provided that the inequality is satisfied, the method applies and we can compute $\alpha=\mathrm{diag}(1,\lambda)$ and $\beta=\alpha\gamma$ from which the Tsirelson inequality (B6) can be determined. Expressed in $\lambda_x^\pm$, one finds that

$$\alpha^*\beta+(\beta^*\alpha)^T=$$

$$\tfrac{2}{\sin\theta_B}\begin{pmatrix}\pm\sin(2x\pm\theta_B) & \mp\sin(2x)\\ \pm(\lambda_x^\pm)^2\sin(2x-\theta_A\pm\theta_B) & \mp(\lambda_x^\pm)^2\sin(2x-\theta_A)\end{pmatrix}$$

and

$$\mathrm{tr}(\alpha^*\alpha+\beta^*\beta)=-\frac{2\sin\theta_A\sin(4x-\theta_A\pm\theta_B)}{\sin(2x-\theta_A)\sin(2x-\theta_A\pm\theta_B)}.$$

Among the possible $\mathbb{P}$ for which the method applies are the probability distributions which lead to maximal violation of a CHSH inequality. The corresponding representations are given by $\theta_A=\theta_B=\pi/2$ and $\psi=\phi_x^\pm$ with $x=\pi/8+n\pi/4$ $(n=0,1,2,3)$.

[1] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, V. Scarani. Phys. Rev. Lett. **98**, 230501. (2007)
[2] E. Hänggi, R. Renner. arXiv:1009.1833 (2010)
[3] L. Masanes, S. Pironio, A. Acín. Nat. Commun. **2**, 238 (2011)
[4] A. Doherty, Y. Liang, B. Toner, S. Wehner. In Proc. of the 23rd Annual IEEE Conference on Computational Complexity, pages 199-210 (2008)
[5] M. Navascues, S. Pironio, A. Acin. New J. Phys. **10**, 073013 (2008)
[6] J. F. Clauser, M.A. Horne, A. Shimony and R. A. Holt. Phys. Rev. Lett. **23**, 880-884 (1969)
[7] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, D. Roberts. Phys. Rev. A **71**, 022101 (2005)
[8] A. Fine. Phys. Rev. Lett. **48**, 291-295 (1982)
[9] `http://qig.itp.uni-hannover.de/qiproblems/1`
[10] L. Masanes. quant-ph/0309137 (2003)
[11] B. S. Tsirelson. J. Soviet Math., **36**(4):557-570, (1985)
[12] R. Impagliazzo, L. Levin and M. Luby. STOC **89**, 12-24 (1989)
[13] W. Arveson. J. Amer. Math. Soc. 21, no. **4**, 1065-1084 (2008)
[14] I. Raeburn, A. M. Sinclair. Math. Scand. **65**, 278-290 (1989)
[15] L. Masanes. quant-ph/0512100 (2005)
[16] N. D. Mermin. Phys. Rev. Lett. **65**, 1838 (1990)
[17] O.Bratteli, D.Robinson. *Operator algebras and quantum statistical mechanics, Band 1.* Springer Verlag (1979)